P63663

D4

PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 4 : H04L 11/20, G06F 15/16

A1

(11) International Publication Number: WO 86/ 02511

(43) International Publication Date: 24 April 1986 (24.04.86)

(21) International Application Number: PCT/US85/01839

(22) International Filing Date: 26 September 1985 (26.09.85)

(31) Priority Application Number: 661,995

(32) Priority Date: 18 October 1984 (18.10.84)

(33) Priority Country: US

(71) Applicant: HUGHES AIRCRAFT COMPANY [US/US]; 200 North Sepulveda Boulevard, El Segundo, CA 90245 (US).

(72) Inventor: McMILLEN, Robert, J. ; 2508 East Willow #104, Long Beach, CA 90806 (US).

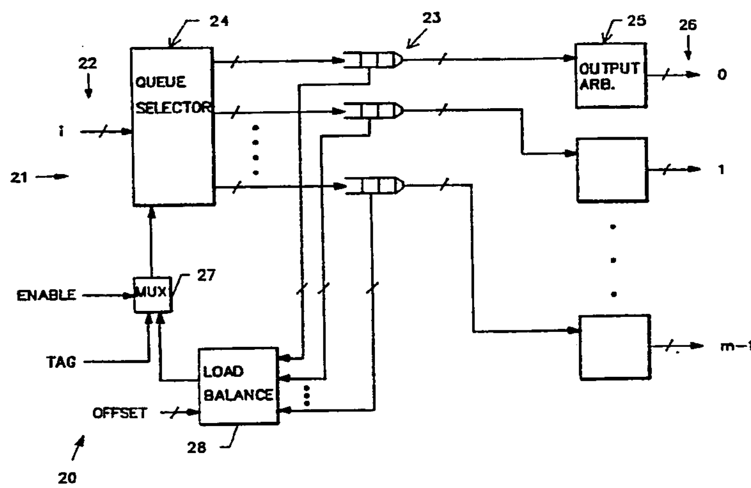
(74) Agents: LINK, Lawrence, V., Jr. et al.; Hughes Aircraft Company, Post Office Box 1042, Bldg. C2, MS A126, El Segundo, CA 90245 (US).

(81) Designated States: DE (European patent), FR (European patent), GB (European patent), IT (European patent), JP, KR, NL (European patent).

Published
With international search report.

(54) Title: LOAD BALANCING FOR PACKET SWITCHING NODES

(57) Abstract



A load balancing circuit arrangement for use with a packet switching node. The packet switching node processes applied data packets containing routing tag signals indicative of the output port destinations to which the data packets are addressed, and routes these packets to the identified output ports. The present invention a load balancing circuit coupled to the packet switching node which monitors the output port addresses of the applied data packets and monitors the number of data packets addressed to each of the output ports. The load balancing circuit is adapted to generate new routing tag signals identifying output port addresses which redistribute the output port load. The load balancing circuit arrangement includes a tag selection circuit coupled to the load balancing circuit and the packet switching node which selectively replaces the routing tag signals of the applied data packets with the new routing tag signals in order to redistribute and balance the output port load. The load balancing circuit comprises a minimum index circuit for generating the new routing tag signals and an adder circuit coupled thereto. The minimum index circuit combines the new routing tag signals with offset signals that modify the new routing tag signals in order to implement a predetermined output port priority scheme. The load balancing circuit arrangement may be employed in both multiple queue and multiport memory packet switching nodes employed in computer or telephone communications applications.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AT	Austria	GA	Gabon	MR	Mauritania
AU	Australia	GB	United Kingdom	MW	Malawi
BB	Barbados	HU	Hungary	NL	Netherlands
BE	Belgium	IT	Italy	NO	Norway
BG	Bulgaria	JP	Japan	RO	Romania
BR	Brazil	KP	Democratic People's Republic of Korea	SD	Sudan
CF	Central African Republic			SE	Sweden
CG	Congo	KR	Republic of Korea	SN	Senegal
CH	Switzerland	LI	Liechtenstein	SU	Soviet Union
CM	Cameroon	LK	Sri Lanka	TD	Chad
DE	Germany, Federal Republic of	LU	Luxembourg	TG	Togo
DK	Denmark	MC	Monaco	US	United States of America
FI	Finland	MG	Madagascar		
FR	France	ML	Mali		

LOAD BALANCING FOR PACKET SWITCHING NODES

BACKGROUND OF THE DISCLOSURE

The present invention relates generally to packet switching nodes employed in multi-processor and parallel computer systems, and the like, and more particularly to a load balancing circuit arrangement for use in such packet switching nodes which redistributes incoming data packets to its output ports for more efficient processing.

One developing area of computer technology involves the design and development of large-scale, multi-processor-based distributed and parallel computer systems. Typical of these classes of computer systems and architectural approaches are the single instruction stream, multiple data stream (SIMD) computer architecture and the multiple instruction stream, multiple data stream (MIMD) computer architecture.

A SIMD computer typically comprises a control unit, N processors, N memory modules and an interconnection network. The control unit broadcasts instructions to all of the processors, and all active processors execute the same instruction at the same time. Each active processor executes the instruction on data in its own associated memory module. The interconnection network provides a communications facility for the processors and memory modules.

A MIMD computer typically comprises N processors and N memories, and each processor can execute an independent instruction stream. Each of the processors may communicate to any other processor. Similar interconnection networks may be employed in the MIMD computer.

Various interconnection networks may be employed to interconnect processors and memories employed in either type of computer system. These interconnection networks include delta networks, omega networks, indirect binary n-cube networks, flip networks, cube networks and banyan networks, for example.

The above-cited networks are discused in some detail in the following publications: "LSI implementation of modular interconnection networks for MIMD machines," 1980 Int'l. Conf. Parallel Processing, Aug. 1980, pp. 161-162; "Analysis and simulation of buffered delta networks," IEEE Trans. Computers, Vol. C-30, pp. 273-282, April 1981; "Processor-memory interconnections for multiprocessors," 6th Annual Int'l. Symp. Computer Architecture, April 1979, pp. 168-177; "Design and implementation of the banyan interconnection network in TRAC," AFIPS 1980 Nat'l. Computer Conf., June 1980, pp. 643-653; "The multistage cube: a versatile interconnection network," Computer, Vol. 14, pp. 65-76, Dec. 1981; "The hybrid cube network," Distributed Data Acquisition, Computing and Control Symp., Dec. 1980, pp. 11-22; and "Performance and implementation of 4x4 switching nodes in an interconnection network for PASM," 1981 Int'l Conf. on Parallel Processing, Aug. 1981, pp. 229-233.

Several types of data switching techniques may be employed to transfer data in SIMD and MIMD computers, and the like, including packet switching, message switching, time-division circuit switching or space-division circuit switching. Packet switching

involves sending one or more words of data at time through the system.

A multiple queue packet switching node is described in a presently copending patent application entitled "Packet Switched Multiple Queue NxM Switch Node and Processing Method," invented by R. J. McMillen, and assigned to the assignee of the present invention. This patent application discloses a packet switching node which processes applied data packets containing routing tag signals indicative of the output port destination to which the data packets are to be applied and transfers these packets to those output ports.

The packet switching node comprises a plurality of input ports and a plurality of output ports. A plurality of queue selectors are individually coupled to corresponding ones of the plurality of input ports. Each of the plurality of queue selectors are adapted to route data packets applied to each of the input ports in accordance with the output port destination of the data packets.

A plurality of queue sets are individually coupled to corresponding ones of the plurality of queue selectors. Each of the queue sets comprise a plurality of queues for storing and forwarding data packets applied thereto as a function of output port destination. A plurality of output arbitrators are individually coupled between corresponding ones of the plurality of output ports and the respective queue of each of the queue sets which store and forward data packets whose destinations are the corresponding output port. The output arbitrators are adapted to transfer the data packets stored in the queues to the corresponding output port in accordance with a predetermined priority arbitration scheme.

Another related presently co-pending patent application is entitled "Packet Switched Multiport Memory NxM Switch Node and Processing Method," invented by R. J. McMillen and A. Rosman, and assigned to the assignee of the present invention. This packet switching node processes applied data packets containing routing tag signals indicative of the output port destination to which the data packets are to be applied. The packet switching node comprises a plurality of input ports and output ports with a multiport memory coupled therebetween. The memory has a predetermined number of memory locations available for storage of data packets applied to each of the input ports. Control logic is coupled to the input and output ports and the multiport memory which controls the storage of data packets in the memory. The control logic also controls the routing of the data packets to the output ports in accordance with the routing tag signals.

In general, this invention comprises an N x M switch node that accepts data packets at any of N input ports and routes each to any of M output ports. The output selected is determined by the routing tag signal in the packet. The control logic is designed so that the data packets are effectively sorted according to their desired output port destination. Arbitration logic randomly, in a statistical sense, chooses among any data packets that are directed to the same output port. The algorithm implemented by the arbitration logic is designed so that data packets will not wait indefinitely to be routed from the switch node.

However, although both of these nodes improve upon the performance and throughput of conventional packet switching designs, their performance can be improved. An example will illustrate a typical problem. Assume that processor 1 applies data packets to

input port 1. Assume that all output ports of the node are connected to identical execution units. Processor 1 assigns routing tag signals to each of the data packets transmitted thereby. If processor 1 continually assigns routing tag signals corresponding to output port 1, for example, then the remaining output ports are not used, and the execution units connected thereto are not used. This clearly lessens the efficiency of the node.

SUMMARY OF THE INVENTION

In order to overcome the limitations of prior art switching node designs, the present invention provides for a load balancing circuit arrangement that improves the efficiency and performance of existing packet switching nodes. In general, the load balancing circuit arrangement of the present invention comprises a logic circuit for use with a packet switching node that processes applied data packets that contain routing tag signals indicative of the output port destinations to which the data packets are addressed.

The load balancing circuit arrangement comprises a load balancing circuit coupled to the packet switching node which monitors the output port addresses of the applied data packets. In addition, the load balancing circuit monitors the number of data packets addressed to each of the output ports. The load balancing circuit is adapted to generate new routing tag signals identifying output port addresses which redistribute the output port load. The load balancing circuit arrangement includes a tag selection circuit coupled to the load balancing circuit and the packet switching node which selectively replaces the routing tag signals of the applied data packets with

the new routing tag signals in order to redistribute and balance the output port load.

One embodiment of the present invention provides for a load balancing circuit arrangement for use with a multiple queue packet switching node that distributes data packets to the least full queues. This operation re-routes the data packets to output ports that are being infrequently used and hence provides for more efficient processing. A second embodiment of the load balancing circuit is provided for use with a multiport memory packet switching node.

In the first embodiment, the switching node comprises queue selection logic coupled to a plurality of queues which route the applied data packets to the desired output port destinations. The load balancing circuit arrangement for use with the multiple queue packet switching embodiment comprises the load balancing circuit coupled between a plurality of queues and their associated queue selection logic. The load balancing circuit monitors the number of data packets stored in each of the queues. The load balancing circuit is coupled to the tag selection circuit which provides input signals to the queue selection logic that replace the routing tag signals of the data packet being processed with new routing tag signals. The new routing tag signals are representative of the queue which has the fewest packets stored therein. This processing redistributes applied data packets to the least full queue and hence redistributes the output port loading of the circuit.

In the second embodiment, the switching node comprises a multiport memory and control logic which similarly route applied data packets to the desired output ports. The load balancing circuit arrangement for use in the multiport memory packet switching

embodiment comprises the load balancing circuit coupled to the control logic which monitors the number and destination of the data packets stored in the multiport memory. The load balancing circuit arrangement provides input signals to the control logic which replace the routing tag signals with new routing tag signals which redistribute the applied data packets to the output ports to achieve balanced output port loading.

The load balancing circuit arrangement comprises a load balancing circuit that monitors the output port addresses of the applied data packets and the number of data packets addressed to each of the output ports. This load balancing circuit generates the new routing tag signals. A tag selection circuit coupled to the load balancing circuit selectively applys the new routing tag signals to the switching node in response to enabling signals applied thereto.

The load balancing circuit comprises a minimum index circuit for generating the new routing tag signals and an adder circuit coupled to the mimimum index circuit for combining the new routing tag signals with offset signals which modify the new routing tag signals in order to implement a predetermined output port priority scheme.

A switching node employing the load balancing circuit arrangement of the present invention may be used to construct a multistage interconnection networks which have several capabilities. The first has the ability to distribute incoming data packets over a pool of shared resources. These resources are typically identical execution or memory units connected to the output ports of the node. The second employs an extra node stage at the input of the network. Enabling the load balancing circuit in the switch nodes of the newly

added stage balances the load throughout the entire multistage network.

Although the present invention is discussed with reference to its use with computer systems and architectures, it is not limited to this application. The present invention may also be used in applications involving the communications field. In particular, interconnection networks employing the present invention may be used to connect telephone systems that communicate both data and voice information by way of data packets.

BRIEF DESCRIPTION OF THE DRAWINGS

The various objects and features of the present invention may be more readily understood with reference to the following detailed description taken in conjunction with the accompanying drawings, wherein like reference numerals designate like structural elements, and in which:

Fig. 1 illustrates a block diagram of an embodiment of a load balancing circuit in accordance with the principles of the present invention incorporated into one channel of a multiple-queue packet switching node;

Fig. 2 illustrates a more detailed block diagram of the load balancing circuit of Fig. 1;

Figs. 3a and 3b comprise a detailed schematic of the logic contained in the minimum index circuit of the load balancing circuit of Fig. 2;

Figs. 4a, 4b and 4c show detailed schematics illustrating the logic contained in the modulo four adder of the load balancing circuit of Fig. 2 for three priority conditions;

Figs. 5a and b comprise a detailed block diagram of a multiport memory packet switching node without load balancing;

Fig. 6 illustrates the load balancing circuit of the present invention coupled to one channel of the multiport memory packet switching node of Fig. 5; and

Fig. 7 illustrates a minimum index circuit for use in the load balancing circuit of Fig. 6.

DESCRIPTION OF THE PREFERRED EMBODIMENTS

Referring to Fig. 1, a block diagram of one embodiment of a load balancing circuit arrangement 20 in accordance with the principles of the present invention is shown. The load balancing circuit arrangement 20 is shown incorporated into one channel of a multiple queue packet switching node 21. This load balancing circuit arrangement 20 is made part of a single stage packet switching node 21. The packet switching node 21 processes applied data packets containing routing tag signals indicative of the output port destination to which the data packets are addressed.

One channel of the packet switching node 21 comprises an input port 22 and a plurality of output ports 26. A queue set 23 is coupled to the input port 22. The queue set 23 comprises a plurality of queues which individually process and store data packets whose destinations are a particular one of the output ports 26. A queue selector 24 is coupled between the input port 22 and each of the plurality of queues in the queue set 23. The queue selector 24 is adapted to route applied data packets arriving at the input port 22 to corresponding ones of the queues in accordance with the addresses contained in the routing tag signals.

An output arbitrator 25 is coupled between each output port and respective queues of the queue set. The output arbitrator 25 processes data packets whose destinations are the output port coupled thereto and applies the data packets stored in each of the respective queues to that output port in accordance with a predetermined priority arbitration scheme.

A more detailed description of the basic multiple queue packet switching node 21 is provided in the presently copending application entitled "Packet Switched Multiple Queue NxM Switch Node and Processing Method," cited hereinabove. This patent application is incorporated herein by this reference.

The load balancing circuit arrangement 20 is coupled between each of the queues of the queue set 23 and the queue selector 24. The load balancing circuit arrangement 20 monitors the number of data packets stored in each of the queues. The load balancing circuit arrangement 20 provides input signals to the queue selector 24 which replace the routing tag signals of the data packets currently being processed. These input signals are representative of the queue in the queue set 23 which has the fewest packets stored therein.

The load balancing circuit arrangement 20 comprises a load balancing circuit 28 coupled to the queues 23 of the packet switching node 20. The load balancing circuit 28 monitors the number of data packets addressed to each of the output ports 26 by examining the number of packets in the queue 23 associated with each output port 26. The load balancing circuit 28 is adapted to generate new routing tag signals identifying output port addresses which redistribute the output port load. A tag selection circuit 27 is coupled between the load balancing circuit 28 and

the queue selector 24 which is adapted to selectively replace the routing tag signals of the applied data packets with the new routing tag signals in order to redistribute and balance the output port load.

The load balancing circuit 28 may comprise a minimum index circuit coupled to an adder circuit. The minimum index circuit is adapted to generate the new routing tag signals and the adder circuit is adapted to combine the new routing tag signals with offset signals which modify the new routing tag signals in order to implement a predetermined output port priority scheme. These two circuits will be described in more detail with reference to Fig. 2. The tag selection circuit 27 may comprise a two to one multiplexer, or the like, which is adapted to selectively choose between two input signals for output to the queue selector 24.

In operation, the multiple queue packet switching node 21 employing the load balancing circuit arrangement 20 functions in the following manner. In a typical multi-processor computer application requiring load balancing, or the like, a process controller, or the like, is coupled to the input port 22. This process controller generates data and instructions which are sent to functionally identical execution units individually coupled to the output ports 26. The data and instructions are sent in data packets which are processed by the packet switching node 21. The switching node 21 provides a communications link between the process controller and the execution units.

Each data packet includes a routing tag signal identifying the particular output port 26, and hence the particular execution unit, to which the packet is to be sent. In addition, the data packet also includes data that is to be operated upon in that execution unit. A data packet enters the switching

node 21 and the queue selector 22 reads the routing tag signal. It then routes the data packet to the appropriate queue of the queue set 23 which processes signals whose destinations are the particular output port 26 corresponding to the routing tag signal.

In the multi-processor computer of this example, each of the output ports are connected to functionally identical execution units. Therefore, it is not particularly important where the instructions contained in the data packets get executed, but that they are executed in a timely manner. Since all of the execution units are functionally identical, routing tag signals may be changed in order to more efficiently utilize the execution units.

The load balancing circuit arrangement 20 produces the desired re-routing of the data packets to accomplish more efficient operation of the switching node 21 and the computer system in general. The load balancing circuit 20 processes count signals generated by each of the queues 23 in the queue set 24. The count signals indicate the number of data packets stored in a particular queue. The load balancing circuit arrangement 20 then analyzes the count signals to determine which queue has the least number of data packets stored therein. Once this has been determined, the routing tag signal in the data packet being processed by the queue selection logic 23 is changed to identify the least full queue.

To better understand the operation of the load balancing circuit 20, reference is made to Fig. 2. Fig. 2 shows a more detailed block diagram of the load balancing circuit 28. The load balancing circuit 28 is comprised of the minimum index circuit 30 and an adder circuit 31. In Fig. 2, the load balancing circuit 28 is shown adapted for use with a node which comprises

four output ports, hence it employs four sets of inputs to the minimum index circuit 30, identfied as L[0] to L[3].

Each set of inputs to the minimum index circuit 30 is individually coupled to count circuitry in the queues which counts the number of data packets currently being stored and processed by a particular queue. The output of the minimum index circuit 30 is coupled to a modulo M adder 31 (modulo 4 in Fig. 2) which adds an offset to the index produced. The load balancing circuit 28 implements an algorithm which determines the least full queue in the queue set. A typical algorithm is presented below.

```
1.   INPUT SCALAR:  OFFSET
2.   INPUT ARRAY:   L[M]
3.   OUTPUT SCALAR: ADDR
4.   BEGIN
5.        minimum:=0;
6.        FOR (i=1; WHILE i<M; i:=i+1)
7.             IF (L[i] < L[minimum])
8.                  minimum:=i;
9.        ENDFOR
10.       ADDR: = (minimum + OFFSET) MODULO M;
11.  END
```

This algorithm scans the inputs L[M] from low order to high, looking for the smallest value. In the case of a tie, the value is retained, thus implementing a particular priority scheme. This priority scheme is as follows. Since there can be a tie between two or more queues for the lowest number, a default priority is assigned. Let L[0] be the first input to the minimum index circuit 30 and let L[M-1] be the last input. In the case of a tie, L[0] has the highest priority and L[M-1] has the lowest.

Alternatively, if the first queue is not to have the highest priority, rather queue i, 0<i<M, then the queue i data packet count is connected to L[0], the queue with the next highest priority is connected to L[1], and so on. In addition, in order to obtain the correct re-routing tag signal, the binary representation of i is connected to the OFFSET input to the modulo M adder 31. To convert this algorithm into a logic circuit requires the specification of the maximum queue size and the number of output ports (M), in order to determine the number of input and output bits in the load balancing circuit 28.

The ability to specify priority is important. In an NxM packet switching node, for example, N of the load balancing circuits are required. If all circuits gave top priority to the lower numbered outputs, there would be a bias in the overall operation of of the switching node favoring these outputs. The distribution of the load over the outputs would not be uniform.

To illustrate how the load balancing circuit 28 is reduced to logic, assume a four by four packet switching node with queues that can hold up to three data packets. Therefore, the number of input ports N = 4 and the number of output ports M = 4, and the number of data packets P = 3. The equation $\lceil \mathrm{Log}_2(P+1) \rceil = \lceil \mathrm{Log}_2(3+1) \rceil = 2$ represents the number of bits required to indicate the data packet count in any queue. Since there are four queues, a total of eight inputs are needed for the minimum index circuit 30. Also, $\lceil \log_2 M \rceil = 2$ bits are required to indicate the which queue is to be selected. Therefore, each input to the minimum index circuit 30, represented by L[i], is represented by two bits, Li1 and Li0, and ADDR and OFFSET are represented by two bits each, ADDR1, ADDR0, and OFFSET1, OFFSET0, respectively.

The load balancing circuit 28 comprises two parts, the minimum index circuit 30 and the modulo four adder 31 which combines the output of the minimum index circuit 30 with the OFFSET value. A modified algorithm is required to implement the minimum index circuit only, and is presented below.

```
INPUT ARRAY:   L[M]
OUTPUT SCALAR: OUT
BEGIN
     minimum:=0;
     FOR (i=1; WHILE i<M; i:=i+1)
     IF (L[i] < L[minimum])
          minimum:=i;
     ENDFOR
     OUT:=minimum;
END
```

This algorithm has been programmed into a computer to generate a truth table with eight boolean input variables and two boolean output variables. This truth table was then made the input to a program which generated the following "sum-of-product" form logic equations. These equations correspond directly to a two level nand gate implementation of the minimum index circuit 30.

$$
\begin{aligned}
OUT1 = {} & \overline{L31}\cdot\overline{L30}\cdot L10\cdot L00 + \overline{L21}\cdot\overline{L20}\cdot L10\cdot L01 \\
& + \overline{L21}\cdot\overline{L20}\cdot L11\cdot L00 + \overline{L31}\cdot L11\cdot L01 \\
& + \overline{L31}\cdot\overline{L30}\cdot L10\cdot L01 + \overline{L31}\cdot\overline{L30}\cdot L11\cdot L00 \\
& + \overline{L21}\cdot\overline{L20}\cdot L10\cdot L00 + \overline{L21}\cdot L11\cdot L01 \\
& + \overline{L30}\cdot L11\cdot L10\cdot L01\cdot L00 \\
& + \overline{L20}\cdot L11\cdot L10\cdot L01\cdot L00 \\
OUT0 = {} & \overline{L11}\cdot\overline{L10}\cdot L00 + \overline{L11}\cdot\overline{L10}\cdot L01 \\
& + \overline{L31}\cdot\overline{L30}\cdot L20\cdot L00 + L20\cdot\overline{L11}\cdot L01 \\
& + \overline{L31}\cdot\overline{L30}\cdot L20\cdot L01 + \overline{L31}\cdot\overline{L30}\cdot L21\cdot L00 \\
& + \overline{L31}\cdot L21\cdot L01 + L21\cdot\overline{L11}\cdot L01 \\
& + L21\cdot\overline{L10}\cdot L01\cdot L00 + \overline{L30}\cdot L21\cdot L20\cdot L01\cdot L00
\end{aligned}
$$

A detailed schematic of the logic diagram representing a specific embodiment of the minimum index circuit 30 is shown in Fig. 3. Fig. 3 is comprised of Figs. 3a and 3b. Figs. 3a and 3b should be placed abutting one another and joined along the line of separation in order to obtain the complete drawing.

The "sum-of-product" form of the equations for the modulo four adder 31 are as follows.

$$C1 = B1\cdot\overline{A1}\cdot\overline{A0} + B1\cdot\overline{B0}\cdot\overline{A1} + \overline{B1}\cdot\overline{B0}\cdot A1 + \overline{B1}\cdot A1\cdot\overline{A0} + \overline{B1}\cdot B0\cdot\overline{A1}\cdot A0 + B1\cdot B0\cdot A1\cdot A0$$

$$C0 = B0\cdot\overline{A0} + \overline{B0}\cdot A0 = B \oplus A$$

Since the B input is the OFFSET constant, once the constant is selected, it is substituted into the above equations which results in simpler circuits. If the OFFSET = 0, no circuit is required. Figs. 4a, 4b and 4c show the circuits to be used when the OFFSET is 1, 2 and 3, respectively. A detailed discussion of the components and wiring of the circuit of Figs. 3 and 4 will be dispensed with.

The load balancing circuit arrangement of the present invention may also be incorporated into a multiport memory packet switching node. Fig. 5 illustrates a portion of a four input, four output multiport memory packet switching node without load balancing. This circuit is identical to the circuit of Fig. 3 of the presently copending multiport memory patent application cited hereinabove. Fig. 5 shows control logic 40 of the multiport memory packet switching node. The control logic comprises four input port logic sections 41a-d coupled to four output port logic sections 42a-d. In addition, handshaking signal lines 45 and status update control logic 46 are included in Fig. 5.

Referring to Fig. 6, there is shown a portion of one channel of the multiport memory packet switching

node of Fig. 5 incorporating a load balancing circuit arrangement 20' of the present invention. The discussion of this embodiment is undertaken with reference to the presently copending patent application entitled "Packet Switched Multiport Memory NxM Switch Node and Processing Method" cited hereinabove with particular reference to Fig. 3 thereof.

Fig. 6 shows a portion of the control logic 40 which includes an input port logic section 41a coupled to four output port logic sections 42a-d. A more detailed understanding of the complete circuit will be had with reference to Fig. 3 in the previously cited patent application. The signal lines illustrated in Fig. 6 of this application correspond to signal lines described in Fig. 3 of the copending patent application.

The load balancing circuit arrangement 20' has inputs coupled to request output lines REQ0, REQ1, REQ2 and REQ3 of the input port logic section 41. Outputs of the load balancing circuit arrangement 20' are coupled to a latch 45 which is used to ensure correct signal timing and then to the enabling circuit 27. A load balancing circuit 28' comprises a minimum index circuit 30' and the adder 31. The minimum index circuit 30' for use in this embodiment is shown in more detail in Fig. 7. The adder circuit 31 is substantially the same as hereinabove described with reference to the load balancing circuit arrangement for the multiple queue node.

Thus, there has been described a load balancing circuit for use with a packet switching node which may be employed in multi-processor and parallel computer systems, and the like. The load balancing circuit may be used to automatically redistribute data packets to external execution units coupled to a multiple queue

or multiport memory switching node. This redistribution results in more efficient operation of the node and computer system employing networks incorporating the load balancing circuit.

It is to be understood that the above-described embodiments are merely illustrative of some of the many specific embodiments which represent applications of the principles of the present invention. Clearly, numerous and varied other arrangements may be readily devised by those skilled in the art without departing from the spirit and scope of the invention. In particular, the present invention may also be used in applications involving the communications field. Interconnection networks employing the principles of the present invention may be used to connect telephone systems that communicate both data and voice information by way of data packets.

CLAIMS

What is claimed is:

1. A load balancing circuit arrangement for use with a packet switching node that processes applied data packets that contain routing tag signals indicative of the output port destinations to which said data packets are addressed, wherein said load balancing arrangement comprises:

load balancing means coupled to said packet switching node for monitoring the output port addresses of said applied data packets and the number of data packets addressed to each of said output ports, and for generating new routing tag signals identifying output port addresses which are adapted to redistribute the output port load; and

tag selection means coupled to said load balancing means and said packet switching node for selectively replacing the routing tag signals of said applied data packets with said new routing tag signals in order to redistribute and balance the output port load.

2. The load balancing circuit arrangement of Claim 1 wherein:

said load balancing means comprises a logic circuit which monitors the output port addresses of said applied data packets and the number of data packets addressed to each of said output ports, and generates said new routing tag signals; and

said tag selection means comprises a logic circuit for selectively applying said new routing tag signals to said switching node in response to enabling signals applied thereto.

3. The load balancing circuit arrangement of Claim 2 wherein said load balancing means comprises:

a minimum index circuit for generating said new routing tag signals; and

an adder circuit coupled to said minimum index circuit for combining said new routing tag signals with offset signals which modify said new routing tag signals in order to implement a predetermined output port priority scheme.

4. A load balancing circuit arrangement for use with a switching node that processes applied data packets containing routing tag signals indicative of the output port destinations to which said data packets are addressed, and wherein said switching node comprises a multiport memory and control logic coupled thereto for routing applied data packets in accordance with the output port destination thereof, wherein said load balancing circuit arrangement comprises:

load balancing means coupled to said control logic for monitoring the output port addresses of said applied data packets and the number of data packets addressed to each of said output ports, and for generating new routing tag signals identifying output port addresses which are adapted to redistribute the output port load; and

tag selection means coupled to said load balancing means and said control logic for selectively replacing the routing tag signals of said applied data packets with said new routing tag signals in order to redistribute and balance the output port load.

5. The load balancing circuit arrangement of Claim 4 wherein:

said load balancing means comprises a logic circuit which monitors the output port addresses of said applied data packets and the number of data packets addressed to each of said output ports, and generates said new routing tag signals; and

said tag selection means comprises a logic circuit for selectively applying said new routing tag signals to said switching node in response to enabling signals applied thereto.

6. The load balancing arrangement of Claim 5 wherein said load balancing means comprises:

a minimum index circuit for generating said new routing tag signals; and

an adder circuit coupled to said mimimum index circuit for combining said new routing tag signals with offset signals which modify said new routing tag signals in order to implement a predetermined output port priority scheme.

7. A load balancing circuit arrangement for use with a packet switching node that processes applied data packets containing routing tag signals indicative of the output port destinations to which said data packets are addressed, and wherein said switching node comprises queue selection logic coupled to a plurality of queues for routing applied data packets in accordance with the output port destination thereof, wherein said load balancing circuit arrangement comprises:

load balancing means coupled to said plurality of queues for monitoring the number of data packets addressed to each of said output ports, and for generating new routing tag signals identifying output port addresses which are adapted to redistribute the output port load; and

tag selection means coupled to said load balancing means and said queue selection logic for selectively replacing the routing tag signals of said applied data packets with said new routing tag signals in order to redistribute and balance the output port load.

8. The load balancing circuit arrangement of Claim 7 wherein:

said load balancing means comprises a logic circuit which monitors the output port addresses of said applied data packets and the number of data packets addressed to each of said output ports, and generates said new routing tag signals; and

said tag selection means comprises a logic circuit for selectively applying said new routing tag signals to said switching node in response to enabling signals applied thereto.

9. The load balancing circuit arrangement of Claim 8 wherein said load balancing means comprises:

a minimum index circuit for generating said new routing tag signals; and

an adder circuit coupled to said minimum index circuit for combining said new routing tag signals with offset signals which modify said new routing tag signals in order to implement a predetermined output port priority scheme.

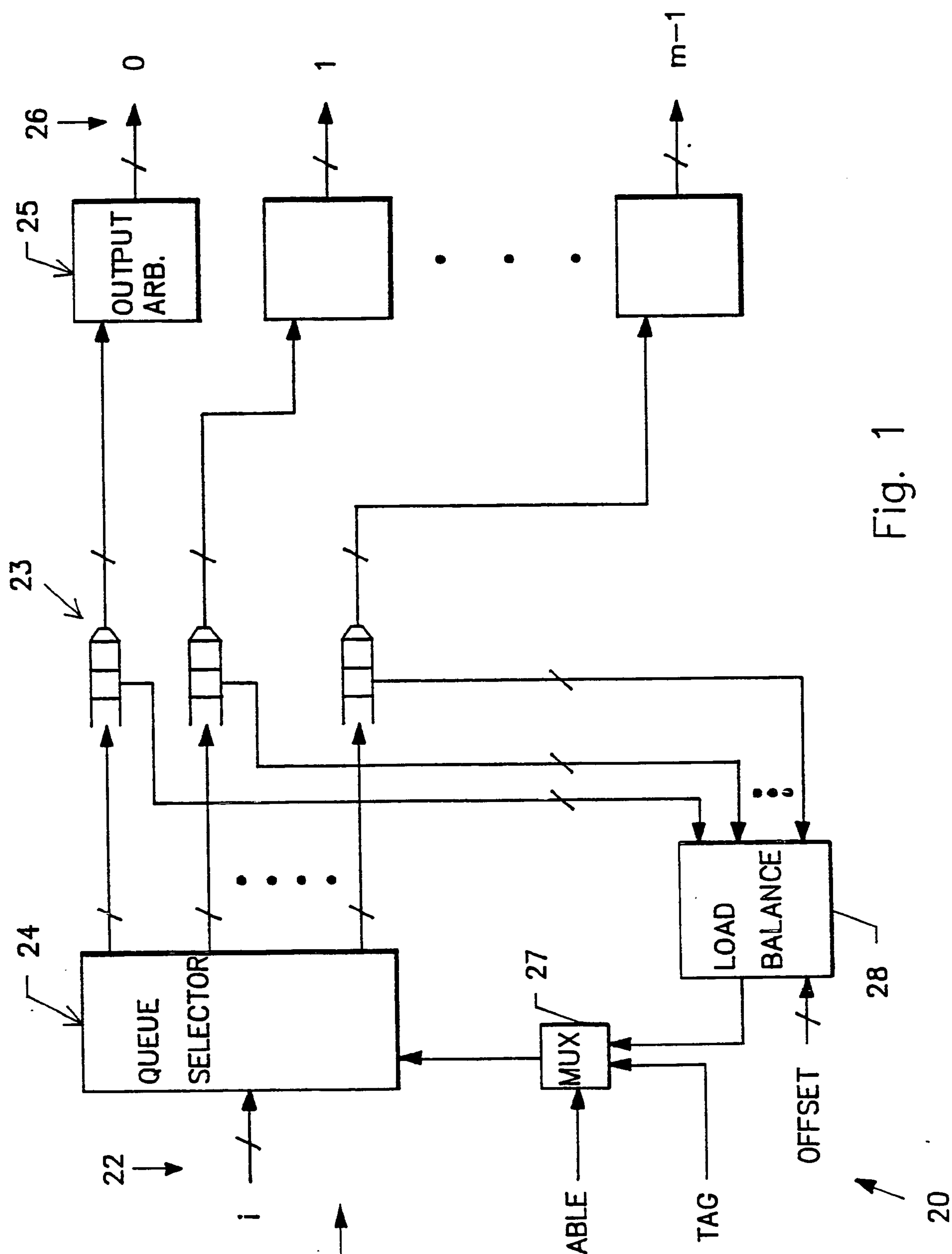


Fig. 1

SUBSTITUTE SHEET

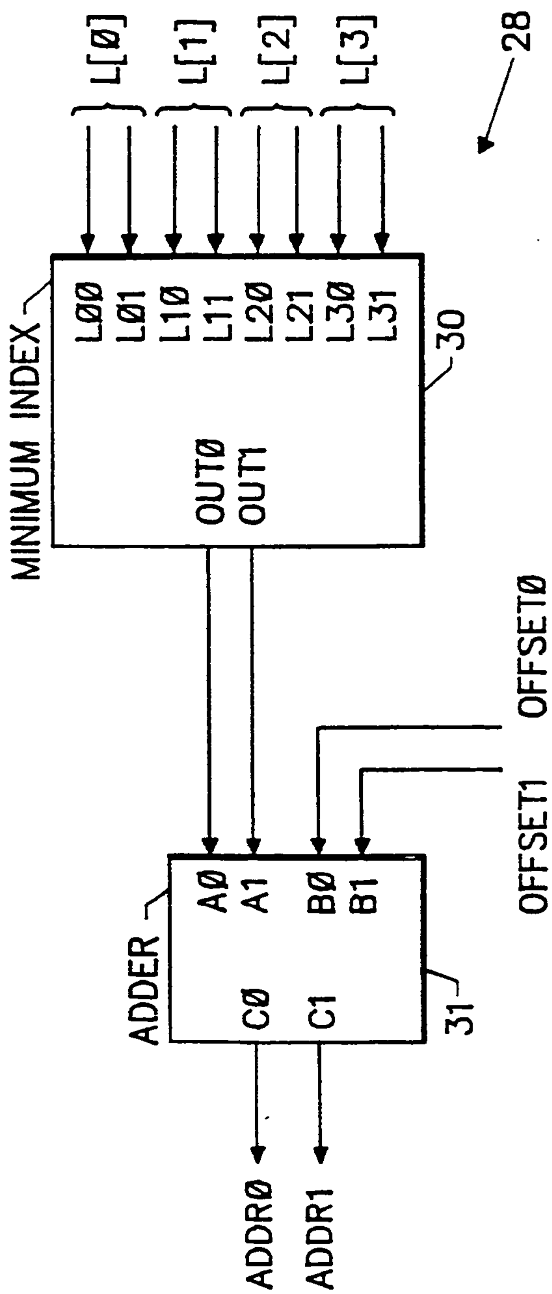

L[Ø]
L[1]
L[2]
L[3]
28
MINIMUM INDEX
LØØ
LØ1
L1Ø
L11
L2Ø
L21
L3Ø
L31
OUTØ
OUT1
30
OFFSETØ
OFFSET1
ADDER
AØ
A1
BØ
B1
CØ
C1
31
ADDRØ
ADDR1

Fig. 2

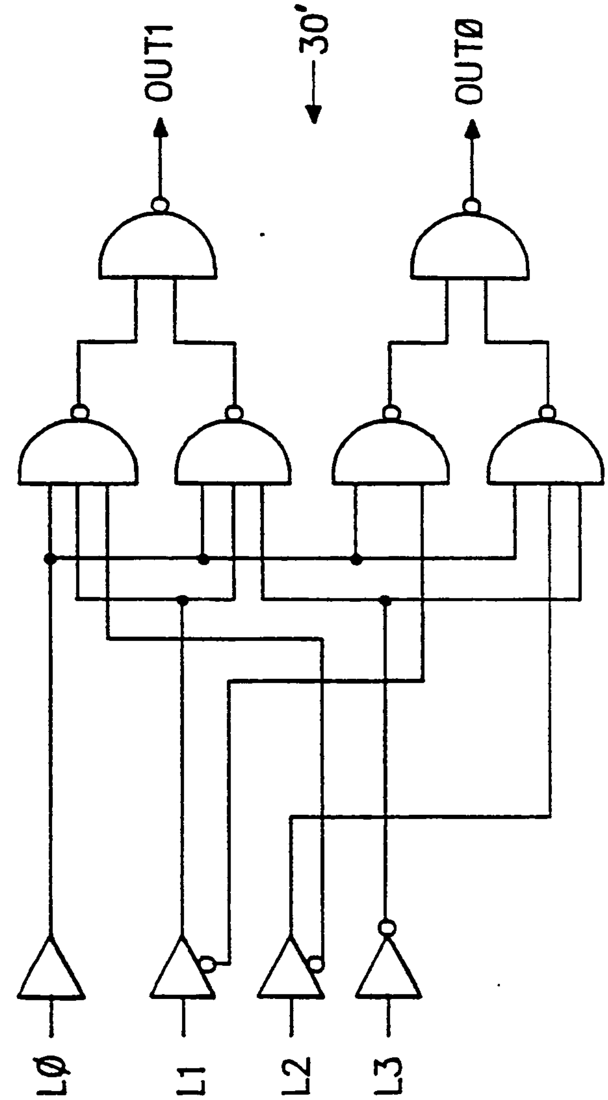
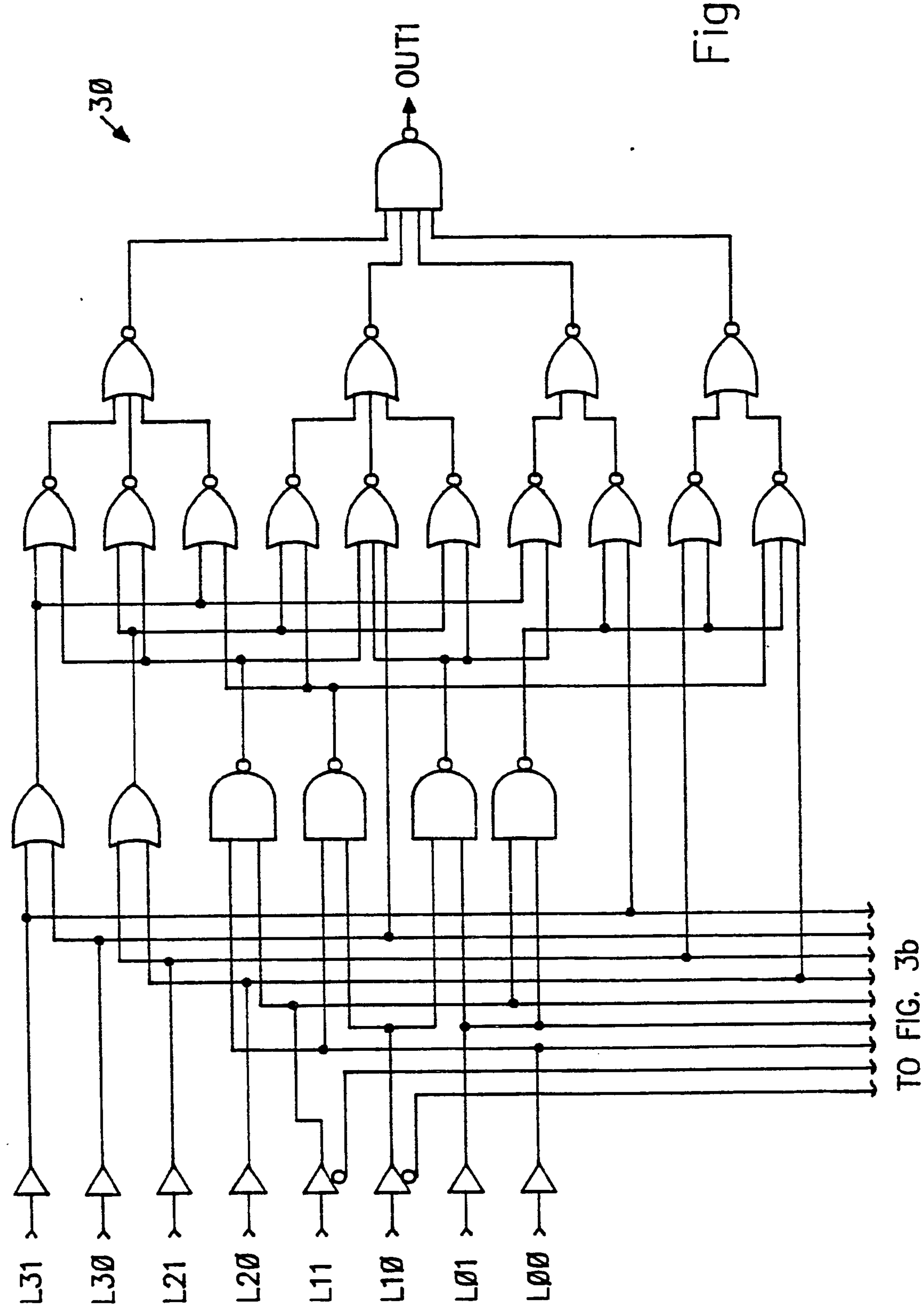

OUT1
30'
OUTØ
LØ
L1
L2
L3

Fig. 7

Fig. 3a



SUBSTITUTE SHEET

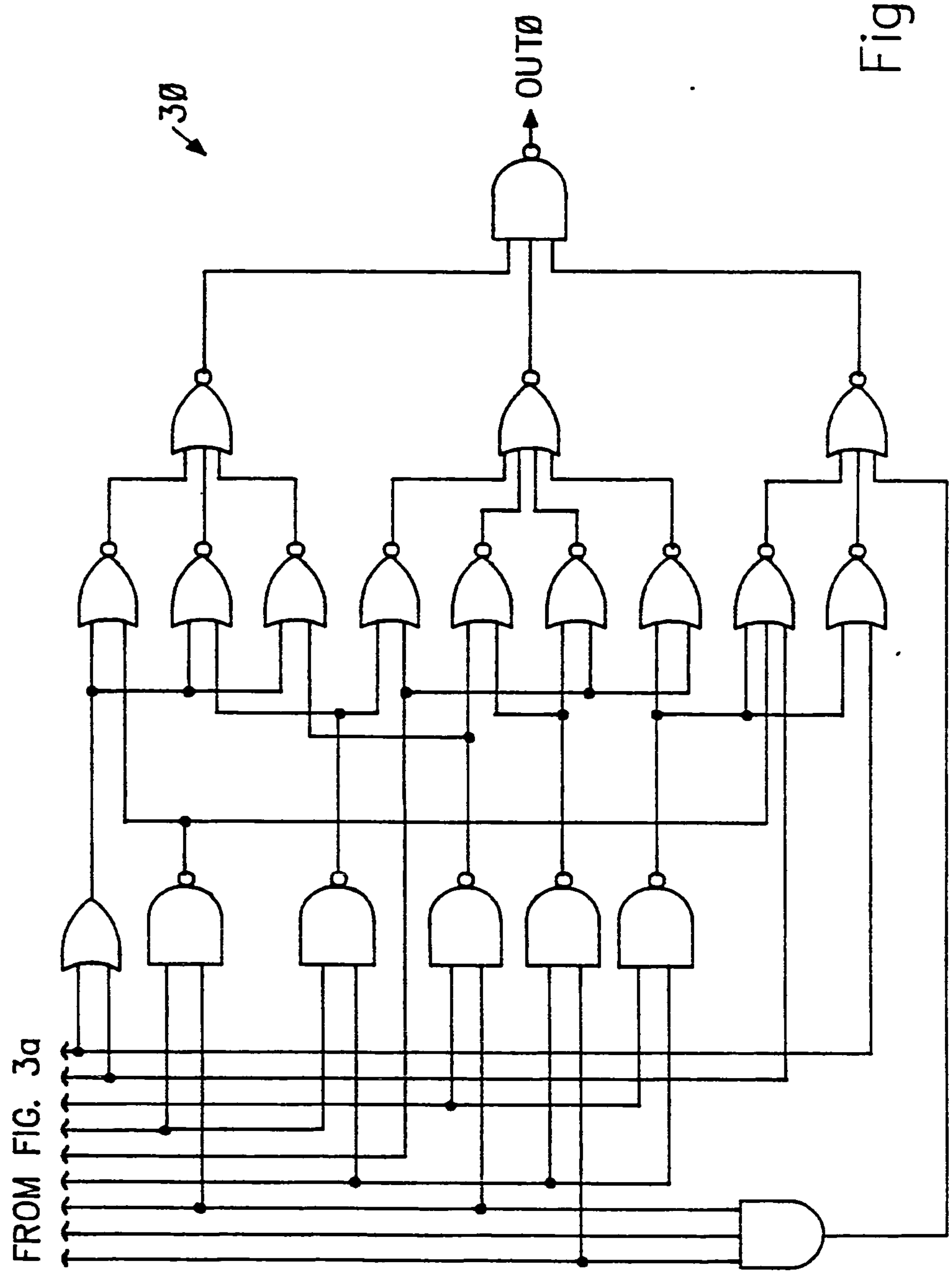


Fig. 3b

SUBSTITUTE SHEET

Fig. 4a

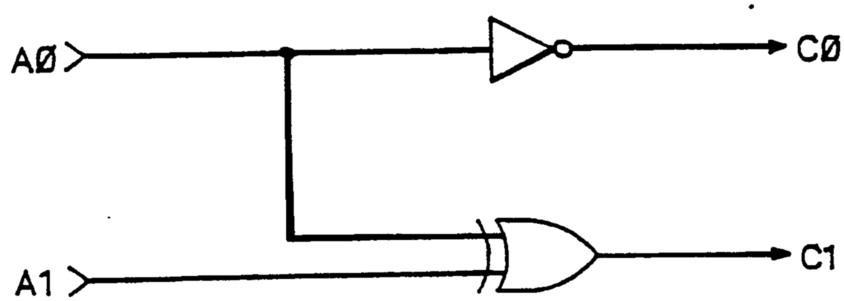


Fig. 4b

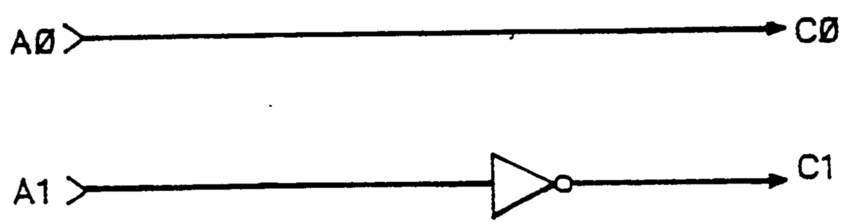
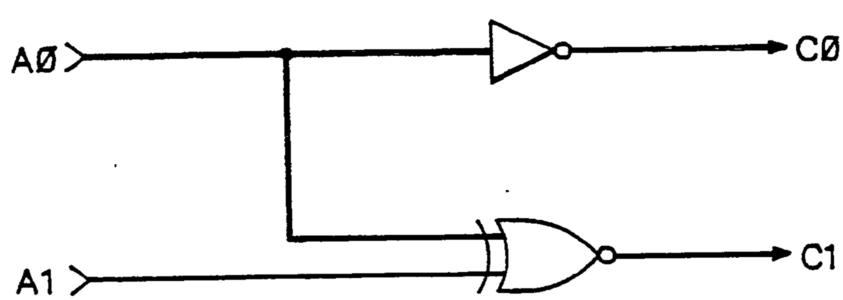


Fig. 4c



SUBSTITUTE SHEET

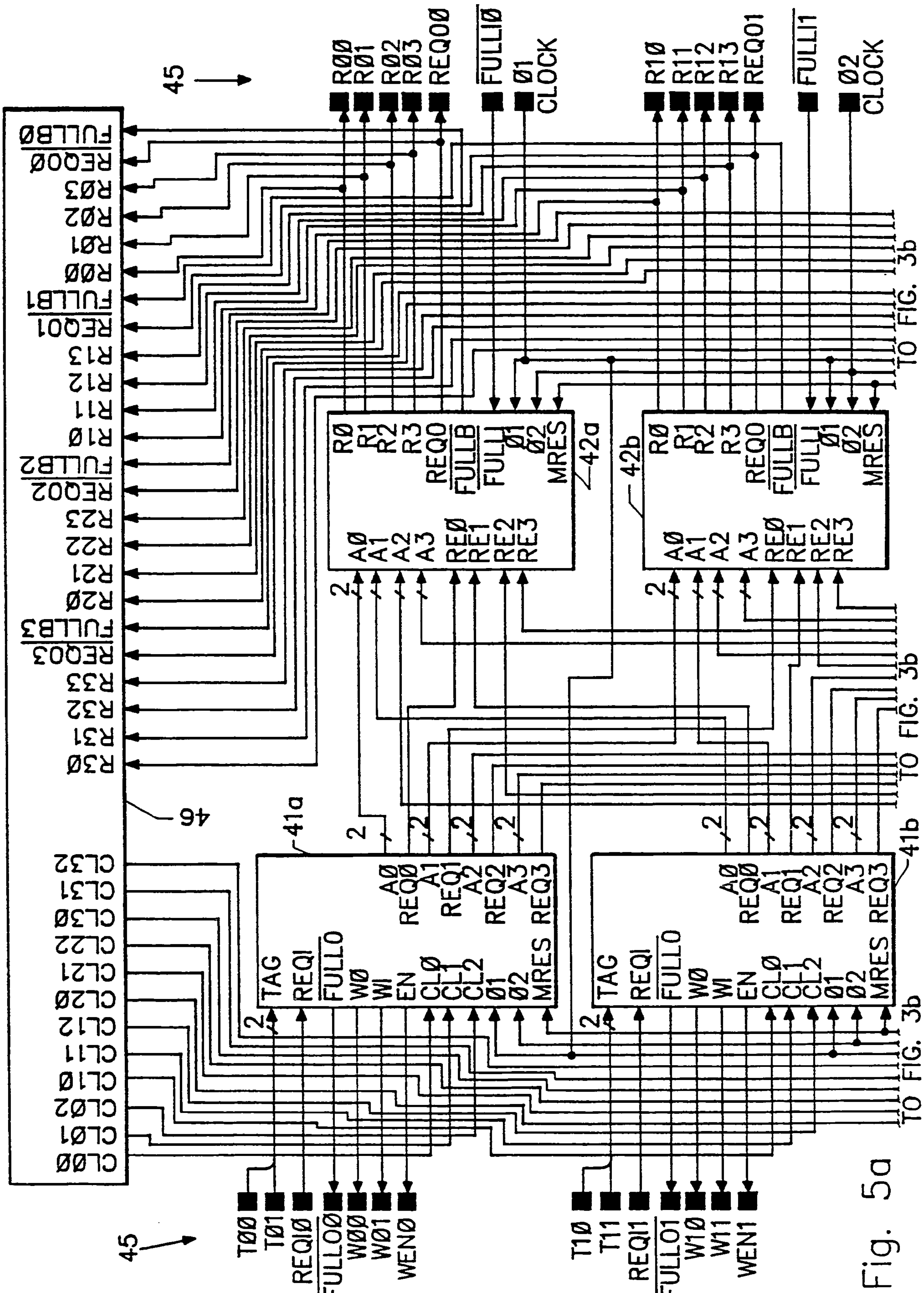

45
R00
R01
R02
R03
REQ00
FULLI0
Ø1 CLOCK
R10
R11
R12
R13
REQ01
FULLI1
Ø2 CLOCK
42a
42b
41a
41b
46
TO FIG. 3b
T00
T01
REQI0
FULLO0
W00
W01
WEN0
T10
T11
REQI1
FULLO1
W10
W11
WEN1

Fig. 5a

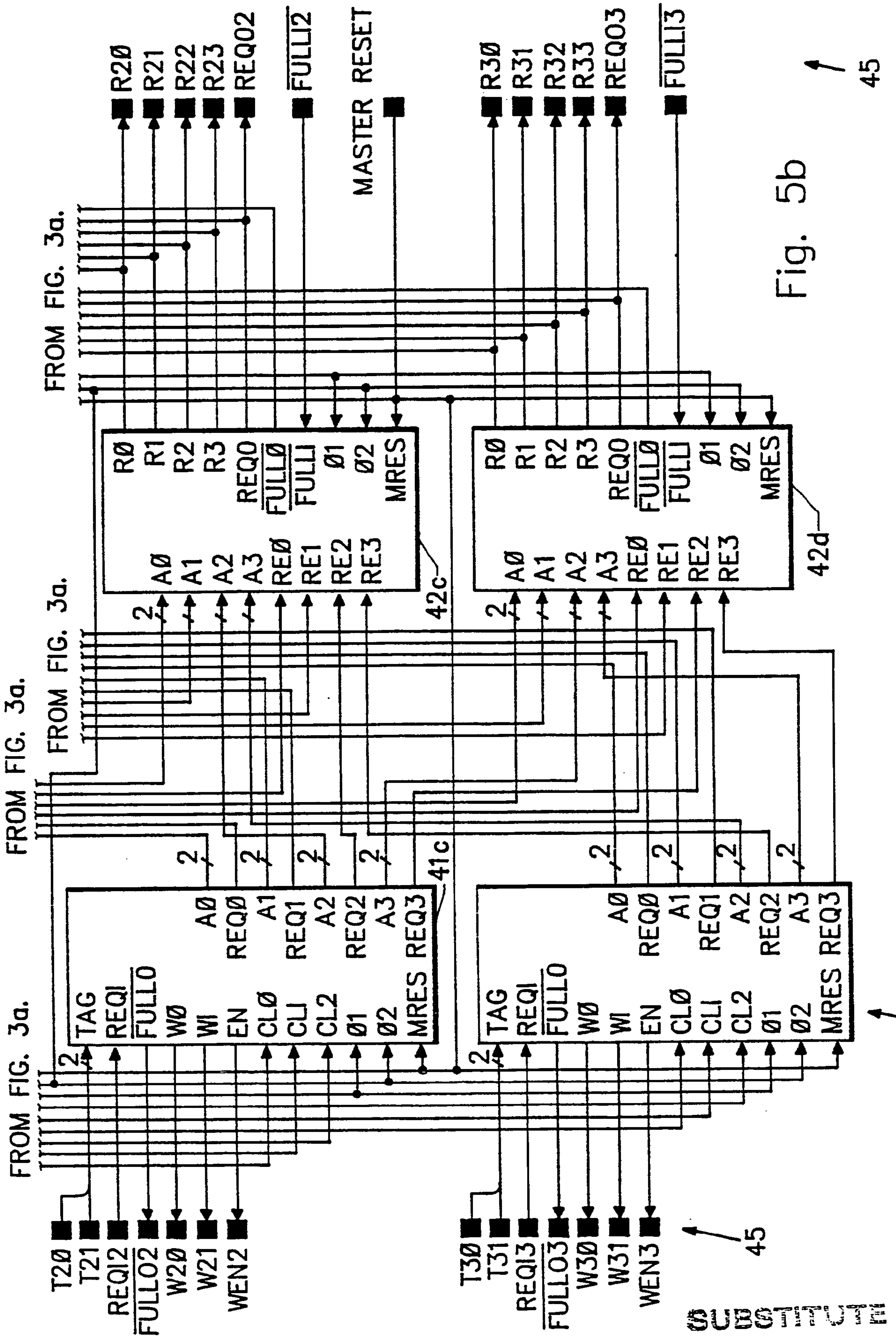


Fig. 5b

SUBSTITUTE SHEET

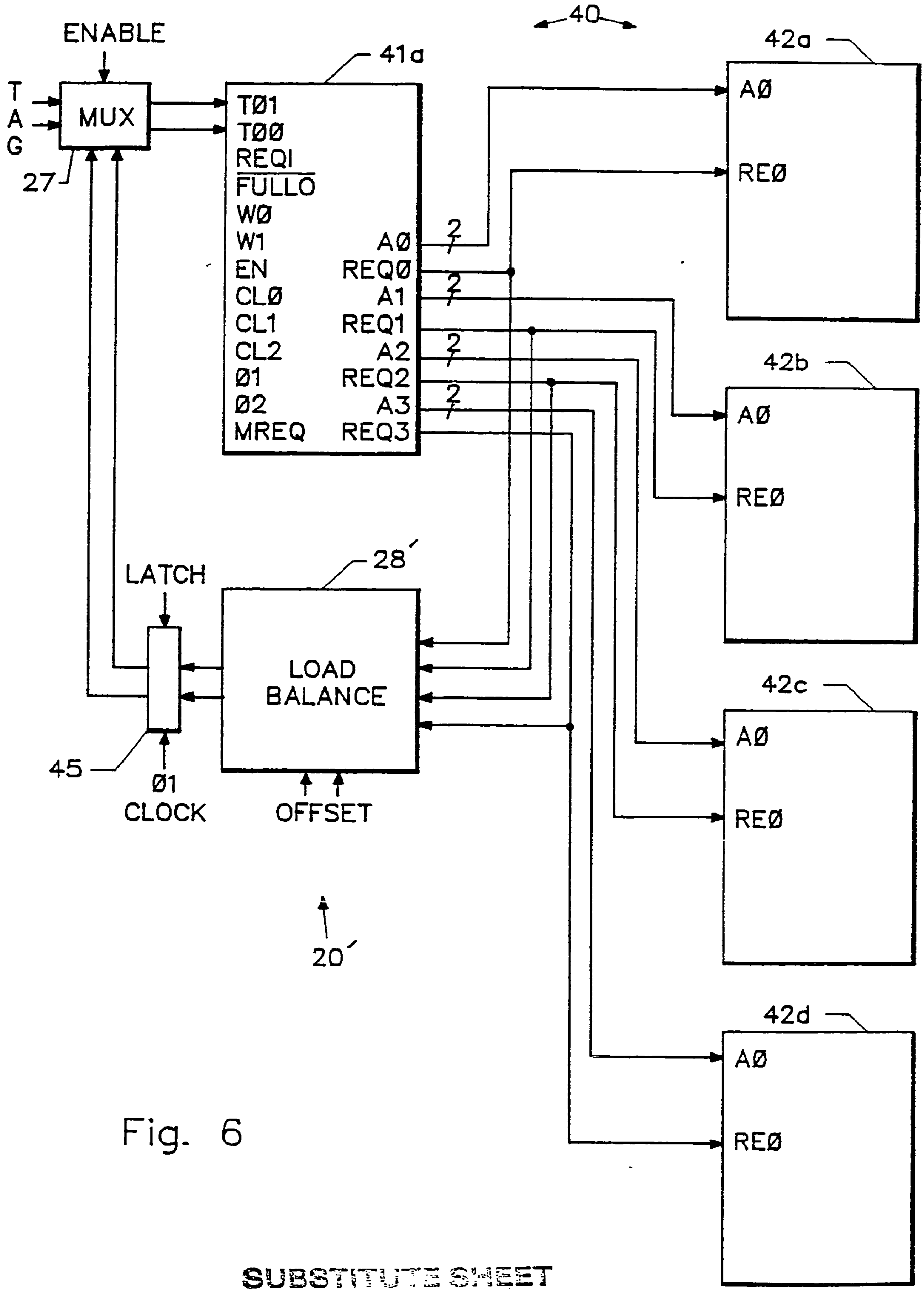


Fig. 6

SUBSTITUTE SHEET

INTERNATIONAL SEARCH REPORT

International Application No PCT/US 85/01839

I. CLASSIFICATION OF SUBJECT MATTER (if several classification symbols apply, indicate all) [8]

According to International Patent Classification (IPC) or to both National Classification and IPC

IPC^4: H 04 L 11/20; G 06 F 15/16

II. FIELDS SEARCHED

Minimum Documentation Searched [7]

Classification System	Classification Symbols
IPC^4	H 04 L 11/20; G 06 F 15/16; G 06 F 13/14

Documentation Searched other than Minimum Documentation to the Extent that such Documents are Included in the Fields Searched [8]

III. DOCUMENTS CONSIDERED TO BE RELEVANT [9]

Category [*]	Citation of Document, [11] with indication, where appropriate, of the relevant passages [12]	Relevant to Claim No. [13]
P,Y	Patents Abstracts of Japan, vol. 8, no. 264 (E-282)(1701), 4 December 1984 & JP, A, 59135953 (FUJITSU) 4 August 1984 see the whole document --	1-9
Y	Proceedings of the International Teletraffic Congress, Stockholm, 13-20 June 1973, Proc. 7, Part 2, Swedish Communications 1973 (Stockholm, SW) Herzog: "Message-switching networks with alternate routing", pages 415/1-415/8, see paragraph 2.1 "Structure and operating made of data switching centers" and figure 2 of page 415/2 --	1-9
Y	WO, A, 84/01077 (WILSON) 15 March 1984 see page 17, lines 1-17 --	1-9
Y	Conference Proceedings: the 9th annual symposium on computer architecture,	./.

* Special categories of cited documents: [10]

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

IV. CERTIFICATION

Date of the Actual Completion of the International Search	Date of Mailing of this International Search Report
19th December 1985	24 JAN. 1986
International Searching Authority	**Signature of Authorized Officer**
EUROPEAN PATENT OFFICE	G.L.M. Kruydenberg

Form PCT/ISA/210 (second sheet) (January 1985)

III. DOCUMENTS CONSIDERED TO BE RELEVANT (CONTINUED FROM THE SECOND SHEET)

Category *	Citation of Document, with indication, where appropriate, of the relevant passages	Relevant to Claim No
	26-29 April 1982 (Austin, US) Parker et al.: "The gamma network: a multiprocessor interconnection network with redundant paths", pages 73-80, see page 74, right-hand column, lines 11-47; figure 1	1-9

ANNEX TO THE INTERNATIONAL SEARCH REPORT ON

INTERNATIONAL APPLICATION NO. PCT/US 85/01839 (SA 10913)

This Annex lists the patent family members relating to the patent documents cited in the above-mentioned international search report. The members are as contained in the European Patent Office EDP file on 17/01/86

The European Patent Office is in no way liable for these particulars which are merely given for the purpose of information.

Patent document cited in search report	Publication date	Patent family member(s)		Publication date
WO-A- 8401077	15/03/84	EP-A-	0104802	04/04/84
		US-A-	4482996	13/11/84

For more details about this annex :
see Official Journal of the European Patent Office, No. 12/82

This Page is inserted by IFW Indexing and Scanning Operations and is not part of the Official Record

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

- [x] BLACK BORDERS
- [x] IMAGE CUT OFF AT TOP, BOTTOM OR SIDES
- [x] FADED TEXT OR DRAWING
- [ ] BLURED OR ILLEGIBLE TEXT OR DRAWING
- [ ] SKEWED/SLANTED IMAGES
- [x] COLORED OR BLACK AND WHITE PHOTOGRAPHS
- [ ] GRAY SCALE DOCUMENTS
- [ ] LINES OR MARKS ON ORIGINAL DOCUMENT
- [ ] REPERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY
- [ ] OTHER: ____________________

IMAGES ARE BEST AVAILABLE COPY.
As rescanning documents *will not* correct images problems checked, please do not report the problems to the IFW Image Problem Mailbox